

KINETIC TOMOGRAPHY I: A METHOD FOR MAPPING THE MILKY WAY'S INTERSTELLAR MEDIUM IN FOUR DIMENSIONS

KIRILL TCHERNYSHYOV J.E.G. PEEK

To appear in The Astronomical Journal (AJ)

Abstract

We have developed a method for deriving the distribution of the Milky Way's interstellar medium as a function of longitude, latitude, distance and line-of-sight velocity. This method takes as input maps of reddening as a function of longitude, latitude, and distance and maps of line emission as a function of longitude, latitude, and line-of-sight velocity. We have applied this method to datasets covering much of the Galactic plane. The output of this method correctly reproduces the line-of-sight velocities of high-mass star forming regions with known distances from Reid et al. (2014) and qualitatively agrees with results from the Milky Way kinematics literature. These maps will be useful for measuring flows of gas around the Milky Way's spiral arms and into and out of giant molecular clouds.

Subject headings: ISM: kinematics and dynamics, methods: statistical

1. INTRODUCTION

Many open problems in star formation, molecular cloud evolution, and galaxy-scale gas dynamics remain open because it has not been possible to measure the most useful quantities for resolving them – the 3D gas velocity vector and 3D gas density over an extended area of sky. A measurement of these fields would allow us to solve the continuity equation (Euler 1757), and derive the rate at which density is changing across the Galaxy over a range of physical scales.

The formation of giant molecular clouds (GMCs), for instance, is in part a matter of collecting a large mass in a small volume. By looking for sites at which gas flows are converging, it may be possible to find currently forming GMCs. Conversely, one could look for diverging flows to detect and characterize the feedback-driven disruption of GMCs. With sufficient spatial and velocity resolution, it would be possible to distinguish between the many theories of exactly how the necessary mass is accumulated and converted to cold, molecular gas. These theories make different predictions for the properties of the required converging flows. For example, some theories assume the converging flows consist mostly of neutral hydrogen (HI) and invoke different sorts of instabilities at the collision interface of the flows to explain how this gas is rapidly converted to cold molecular hydrogen (H₂) (e.g. Heitsch et al. 2006; Clark et al. 2012; Carroll-Nellenback et al. 2014). Other theories assume that GMCs form from the collision and agglomeration of smaller molecular cloudlets (e.g. Roberts & Stewart 1987; Dobbs 2008; Tasker & Tan 2009). One could directly distinguish between these two possible modes of GMC formation by determining whether gas that is converging on forming GMCs is predominantly neutral or molecular. Without the ability to measure the velocity field as a function of all three spatial dimensions, it is difficult to even determine where converging flows are present.

The origin of the converging flows invoked above is also a matter of active interest. In some theories, the self-gravity of a modest overdensity can be sufficient to

induce collapse (Kim & Ostriker 2002; Vazquez Semadeni et al. 2007; Dobbs et al. 2012). Others invoke spatially coherent flows driven by feedback from star formation (Fujimoto et al. 2014) or perturbations in the Galactic potential such as spiral arms (Roberts 1972; Bonnell et al. 2006). Strong, shocked flows driven by spiral arms have been seen in strongly tidally interacting two-arm spiral galaxies in the nearby universe (Visser 1980a,b; Shetty et al. 2007) using CO and HI observations. Unfortunately, the resolution in HI required to map these spiral shocks beyond very nearby galaxies with extreme two-armed spiral structure, is not observationally feasible (Visser 1980a).

The nearest spiral galaxy is, of course, our own Milky Way. Studying the kinematics of the Milky Way replaces the problem of insufficient sensitivity and spatial resolution with the problem of confusion — from our vantage point, it is difficult to determine how the ISM is moving as a function of 3D position. In particular, it is essentially impossible to obtain the transverse velocity field of the ISM. Transverse velocities can only be derived from proper motions, which are difficult to measure for the diffuse and continuous ISM. For many of the open problems we have outlined, even a measurement of the line-of-sight (radial) velocity as a function of 3D position (the 1-velocity field) would represent a significant step forward. While only having one component of the velocity field does make it difficult to empirically measure the rate of inflow and outflow of matter relative to any given structure, it should still be possible to empirically measure this rate statistically across a sample of structures.

No single observable can simultaneously measure the density and line-of-sight velocity at each location in the ISM, a construct we call the position-position-distance-velocity (PPDV) 4-cube. Instead, we have a number of observables of the ISM, each of which can be considered a (noisy and biased) projection of this PPDV 4-cube. We dub the process of reconstructing this PPDV 4-cube from some set of observations “Kinetic Tomography” (KT). There are a number of existing kinds of data we can use. The first are classical radio observations of

the ISM in both diffuse tracers (e.g. HI) and denser tracers (e.g. CO). These are position-position-velocity (PPV) 3-cubes, a specific projection of the PPDV 4-cube, and have been classically used to infer the density of the ISM in 3-space assuming a Galactic rotation curve (e.g. Levine 2006, and references therein). Another observable is the position-position-distance (PPD) reddening 3-cubes generated by examining the photometry of large numbers of stars and performing inference on the intervening dusty ISM. There has been dramatic progress in this field (Marshall et al. 2006; Lallement et al. 2014; Green et al. 2015), which has been crucial for allowing this investigation. Interstellar absorption lines toward stars also represent a projection of the PPDV 4-cube and can simultaneously contain distance, column density, and velocity information about the intervening matter (Welsh et al. 2010; Zasowski et al. 2015; Zasowski 2015).

There have been some attempts to construct maps or point estimates of v_{los} as a function of distance. Most such attempts have focused on individual spiral arms and used models of ISM flows around the spiral arms to directly invert PPV 3-cubes (e.g. Shane 1972; Foster & MacWilliams 2006). Reid et al. (2016) have developed a different approach, which combines probability distributions from the standard kinematic distance, various geometric hints, and possible associations of emitting gas with structures that have parallax-based distance measurements into a combined syncretic probability distribution for the distance. Neither of these approaches uses the information available in reddening-based PPD 3-cubes.

In the method we describe in this work, we use large-area CO and HI PPV 3-cubes and the Green et al. (2015) (henceforth GSF) PPD 3-cube to reconstruct the ISM PPDV 4-cube. We perform a restricted version of the full tomographic reconstruction in which we assume each parcel of gas in the PPD 3-cube is assigned a single central line-of-sight velocity with some line-of-sight velocity width. This can be considered an inversion of the usual kinematic distance method, in which a line-of-sight velocity is converted to a distance using a Galactic rotation curve. Here we map distance along a sightline to velocity and allow deviations from a rotation curve in order to better match the distribution of matter in the PPV 3-cube.

In this work, we describe our method, the map it produces, and our evaluation of this map’s accuracy and precision. Subsequent papers in this series will analyze this map to gain insights into Milky Way kinematics and their connection to star formation. In Section 2, we describe the datasets we use to make and evaluate our PPDV map. In Section 3, we give a detailed explanation of the PPDV mapping technique and quantitatively demonstrate the accuracy of the technique’s results. In Section 4, we discuss the broader accuracy and applicability of the technique. We conclude in Section 5.

2. DATA

2.1. HI and CO data

Radio emission lines of HI and CO trace the two dominant constituents of the Galactic ISM, atomic and molecular gas. Ionized phases of the ISM do not contribute significantly to the column density and will therefore con-

tribute negligibly to the extinction measured in GSF. 21-cm line emission from the hyperfine transition of HI is usually optically thin and its integral is an excellent tracer of HI column:

$$N(\text{HI}) = 1.8 \times 10^{18} \text{ cm}^{-2} \frac{\int T_B dv}{\text{K km s}^{-1}}. \quad (1)$$

When the 21-cm line becomes optically thick, Equation 1 will underestimate the HI column. However, this mostly happens in H_2 -dominated regions (Goldsmith et al. 2007).

We trace molecular gas using the 115 GHz 1-0 rotational transition of CO. The integral of this emission line can be converted to a H_2 column density using the conversion factor (Bolatto et al. 2013)

$$X_{\text{CO}} = 2.0 \times 10^{20} \text{ cm}^{-2} \frac{\int T_B dv}{\text{K km s}^{-1}}. \quad (2)$$

This conversion factor has a number of known weaknesses stemming from complex excitation and opacity effects and real variation in the relative population of CO and H_2 molecules. We will address the impacts of these weaknesses in §4.2.

For our CO data, we use the interpolated whole-Galaxy PPV 3-cube provided by Dame et al. (2001). The 3-cube covers the full range in ℓ , -30° to $+30^\circ$ in b , and -320 km s^{-1} to $+320 \text{ km s}^{-1}$ in V_{LSR} and has a velocity resolution of 1.3 km s^{-1} . We find that the native resolution and PPV extent of these data are appropriate for our investigation and retain their exact pixelization for our other datasets.

The CO 3-cube contains single-pixel artifacts in both emission and absorption. To remove these artifacts, we apply a plus-shaped median smoothing kernel to each velocity channel. This kernel is three pixels wide along the ℓ and b directions. This filtering procedure changes the total amount of CO emission by about 5% over the entire Galaxy.

For HI data, we use a combination of three large-area Galactic HI surveys. South of declination 0° , we use data from the $16'$ resolution GASS survey (Kalberla et al. 2010); from declination 0° to 38° we use unpublished data from the $4'$ resolution GALFA-HI survey (Peek et al. 2011) Data Release 2; North of 38° we default to the $36'$ resolution LAB Survey (Kalberla et al. 2005). We regrid these data onto the $7.5' \times 7.5' \times 1.3 \text{ km s}^{-1}$ pixels of the Dame et al. (2001) CO 3-cube. We note that each of these surveys has potential pitfalls. LAB is quite low resolution compared to the other data sets, meaning small scale Galactic features may be lost. GALFA-HI does not have stray radiation correction applied and so may overestimate columns in low column density regimes. Small artifacts exist in all of these databases but are typically most pronounced in high-latitude fields. The possibility of differences in the accuracy of the KT solution between regions in which we use different HI surveys is discussed in §3.3.

The HI and CO emission 3-cubes are converted to $N(\text{HI})$ and $N(\text{H}_2)$ 3-cubes using Equations 1 and 2. These two column density 3-cubes are then added to make a single N_{H} data 3-cube.

2.2. Dust data

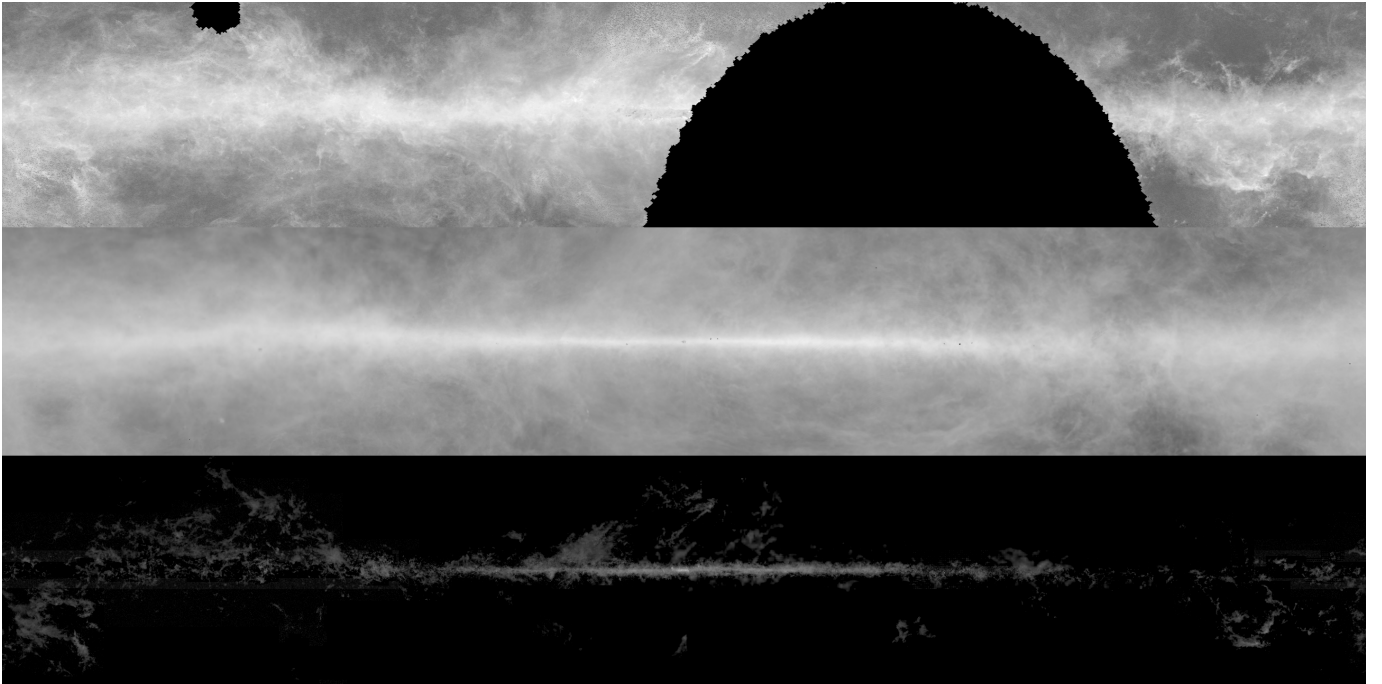


FIG. 1.— The interstellar medium in the Galactic plane, shown in three tracers: reddening (top), HI (middle), and CO (bottom). Each shows the total column measured by each tracer, with a log intensity stretch. The regions are 360 degrees in Galactic longitude wide, centered on Galactic Center, and 60 degrees in Galactic latitude tall, centered on the Galactic plane.

Our extinction 3-cube is derived from the GSF reddening data. GSF use PanSTARRS photometry of 800 million stars to infer the cumulative reddening along the line of sight in $6.8'$ (NSIDE=512 HEALPix) pixels. The distance axis of the GSF 3-cube is in steps of half a distance modulus, from 63 pc to 63 kpc. We regrid these data onto the Dame et al. (2001) ℓ - b grid and difference them in distance to find the reddening between each distance step. This differential reddening is then converted to an N_H using the factor measured in Peek (2013),

$$N_H = E(B - V) 7 \times 10^{21} \frac{\text{cm}^{-2}}{\text{mag}}. \quad (3)$$

We use a single reddening to hydrogen column density conversion factor. This is, in principle, incorrect due to variations in both the relation between the amount of reddening and the amount of dust (e.g. Planck Collaboration and Fermi Collaboration et al. 2015) and the ratio of dust to gas (e.g. Savage & Sembach 1996; Jenkins 2009). We expect the latter to be the more important effect. The combination of HI and CO emission lines traces gas from at least three of the four cloud types in Savage & Sembach (1996), all of which have different depletion patterns and hence dust-to-gas ratios. This change in depletion pattern corresponds to a change in dust-to-gas ratio of about two, which while not negligible is also unlikely to be the dominant systematic uncertainty in our analysis. We discuss the impact of making this incorrect assumption in §4.2.

2.3. High-mass star forming region data

To check the accuracy of our method, we need measurements of the line-of-sight velocities and distances of clouds of gas. Reid et al. (2014) (henceforth Reid14) have

measured the line-of-sight velocities, proper motions, and trigonometric parallaxes of water and methanol masers associated with 103 high-mass star forming regions (HMSFRs). Of these 103, 99 fall inside the footprint of the GSF reddening data. We adopt the line-of-sight velocities, velocity uncertainties, and parallaxes of these 99 HMSFRs as stated in Table 1 of Reid14.

3. KINETIC TOMOGRAPHY

We have developed a procedure for deriving the distribution of interstellar matter in PPDV space from measurements of its distribution in PPD and PPV space. The technical term for deriving a multi-dimensional distribution from lower-dimensional measurements is *tomographic reconstruction*; in this context, lower-dimensional measurements are called *projections*. Tomographic reconstruction from two projections is, in general, not possible – there are many more independent variables than observational constraints. Our procedure for solving this specific case of the tomographic reconstruction problem uses simplifying assumptions about the structure of the ISM in PPDV space to reduce the number of independent variables.

For motivation, we first examine the assumptions behind a common technique for reconstructing a PPDV-space ISM distribution from PPV-space measurements alone. This is the widely known kinematic distance method, which maps line-of-sight velocities to distances based on an assumed rotation curve and Galactic geometry. While the kinematic distance method is usually presented as a way of converting a PPV-space distribution to a PPD-space, rather than PPDV-space, distribution, the two conversions are equivalent according to the kinematic distance method’s underlying assumptions. These assumptions can be combined into a single statement: (1)

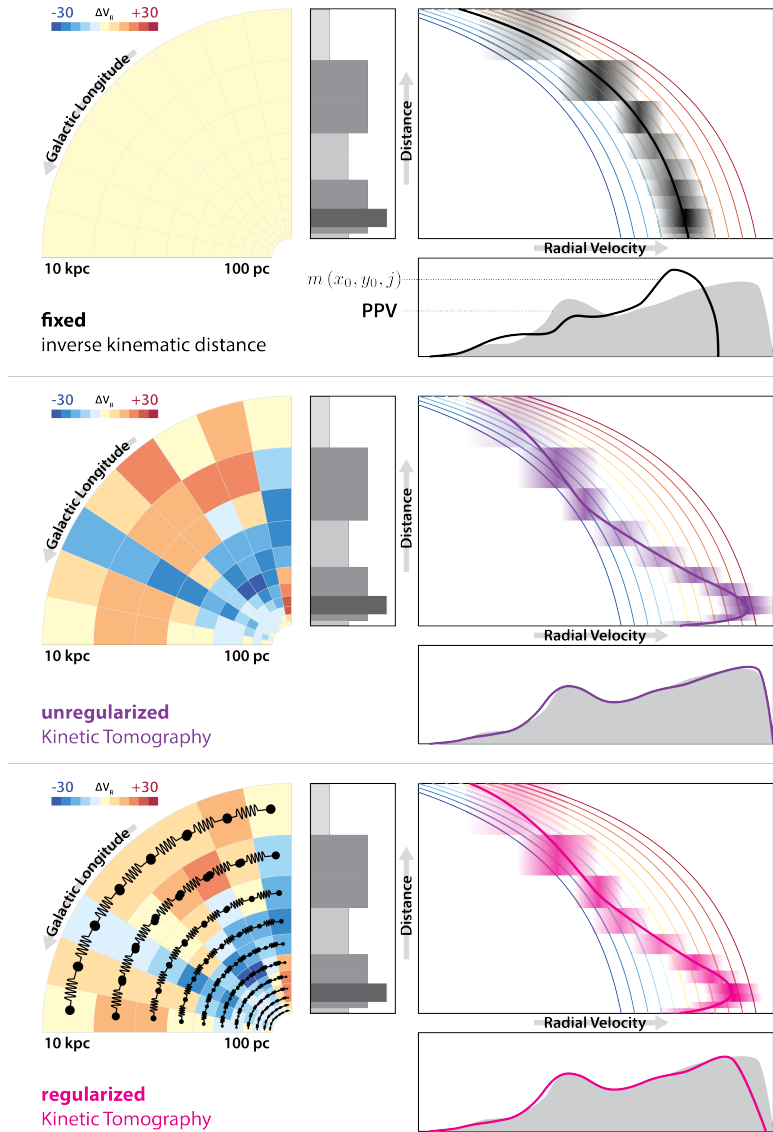


FIG. 2.— Diagrams explaining the inverse kinematic distance methods, unregularized Kinetic Tomography, and regularized Kinetic Tomography. In each diagram, we show a map of residuals from a flat rotation curve (left quarter-circle), mock PPD and PPV data along a single sightline (x_0, y_0) (narrow vertical and horizontal panels), and the mean $\langle v_{\text{los}}(x_0, y_0, i, j) \rangle$ and standard deviation $\langle \sigma_v(x_0, y_0, i, j) \rangle$ of the velocity along that sightline. The ℓ , b , distance, and v_{los} axes are indexed by x, y, i , and j ; note that the residual map shows only a single latitude. The results of applying a fixed inverse kinematic distance method, unregularized Kinematic Tomography, and regularized Kinetic Tomography are shown in black, purple, and pink. Regularization is indicated in the bottom schematic with spring symbols. Our regularization method minimizes the velocity difference between voxels that are connected by spring symbols.

a location in PPD space can be assigned a single line-of-sight velocity (2) according to an assumed rotation curve and Galactic geometry. With these assumptions, it follows that knowing the PPD-space distribution of the ISM is equivalent to knowing its PPDV-space distribution.

We use PPD measurements in addition to PPV measurements, allowing us to relax these assumptions. Our version: (1) a parcel of interstellar matter in PPD space can be assigned a *Gaussian distribution* of line-of-sight velocities (2) whose center is *within a fixed range* of the line-of-sight velocity predicted by an assumed rotation curve and Galactic geometry. Here, a “parcel” of interstellar matter refers to the contents of a single PPD 3-cube voxel. That is to say, we aim to assign a central line-of-sight velocity v_{los} and a line-of-sight Gaussian ve-

locity width σ_v to each voxel in the PPD 3-cube. From our assumptions, a description of the ISM in PPDV space consists of a description of its PPD-space distribution (the observed PPD 3-cube), a line-of-sight central velocity 3-cube in PPD space $(v_{\text{los}}(\ell, b, d))$, and a line-of-sight velocity width 3-cube in PPD space $(\sigma_v(\ell, b, d))$. Thus, we have reduced our original problem of finding a PPDV 4-cube which is consistent with our observed PPD and PPV 3-cubes to finding a $v_{\text{los}}(\ell, b, d)$ and $\sigma_v(\ell, b, d)$ pair consistent with our PPV observations.

3.1. Formalism

We are looking for a $v_{\text{los}}(\ell, b, d)$ and $\sigma_v(\ell, b, d)$ pair that, when combined with the observed PPD 3-cube, best matches the observed PPV 3-cube. This is an opti-

mization problem. An optimization problem needs an objective function, and an objective function needs a quantity to compare to the data. In our case, that quantity is a model PPV 3-cube. To obtain a model PPV 3-cube, we produce a PPDV 4-cube and then integrate it along the distance axis.

Suppose that we have a matter density $\rho(\ell, b, d)$ in PPD space and a $v_{\text{los}}(\ell, b, d)$ and $\sigma_v(\ell, b, d)$ pair. The matter density at a point (ℓ, b, d, v) in PPDV space is

$$\rho(\ell, b, d, v) = \frac{\rho(\ell, b, d)}{\sqrt{2\pi\sigma_v(\ell, b, d)^2}} \exp\left(-\frac{(v - v_{\text{los}}(\ell, b, d))^2}{2\sigma_v(\ell, b, d)^2}\right), \quad (4)$$

and the matter density at a point (ℓ, b, v) in PPV space is

$$\rho(\ell, b, v) = \int_0^{d_{\text{max}}} \rho(\ell, b, d, v) dd, \quad (5)$$

where d_{max} is the maximum distance to which $\rho(\ell, b, d)$ is known.

These equations apply for continuous matter density and v_{los} fields. In our case, all quantities are defined on discrete grids. Instead of a PPD-space matter density $\rho(\ell, b, d)$, for example, we have a PPD 3-cube whose entries are total masses of protons $m(x, y, i)$ inside a voxel centered on ℓ_x, b_y , and d_i .

To move from the continuous case to the discrete case, we assume that $v_{\text{los}}(\ell, b, d)$ and $\sigma_v(\ell, b, d)$ are constant across a voxel and work with integrated quantities instead of densities. Equation 4 becomes

$$m(x, y, i, j) = m(x, y, i) \times \int_{v_j - \Delta v}^{v_j + \Delta v} \frac{1}{\sqrt{2\pi\sigma_v(x, y, j)^2}} \exp\left(-\frac{(v - v_{\text{los}}(x, y, j))^2}{2\sigma_v(x, y, j)^2}\right) dv, \quad (6)$$

where $2\Delta v$ is the length of a voxel along the velocity axis. Equation 5 becomes

$$m(x, y, j) = \sum_i m(x, y, i, j). \quad (7)$$

A cartoon representation of this procedure is shown in the top panel of Figure 2. In the cartoon, $v_{\text{los}}(x, y, i)$ is assumed to be set by a rotation curve; the resulting $m(x, y, j)$ does not match the cartoon PPV data. This mismatch is also generally true of the actual data. Assuming a rotation curve and propagating it through produces a model PPV 3-cube which is clearly inconsistent with the observed PPV 3-cube.

We quantify the discrepancy between a model PPV 3-cube, $m(x, y, j)$, and the observed PPV 3-cube, $PPV(x, y, j)$, with the objective function

$$\mathcal{L}_u(v_{\text{los}}(x, y, i), \sigma_v(x, y, i)) = \frac{1}{2} \sum_{x, y, j} (PPV(x, y, j) - m(x, y, j))^2, \quad (8)$$

the sum of square differences between the model and the observations. This is the *unregularized* objective function; hence the subscript u in \mathcal{L}_u .

We optimize \mathcal{L}_u by varying the entries of $v_{\text{los}}(x, y, i)$ and $\sigma_v(x, y, i)$. We restrict $\sigma_v(x, y, i)$ to be between

1 and 15 km sec⁻¹. The lower bound on $\sigma_v(x, y, i)$ is set to be slightly smaller than $2\Delta v$, which in this case is 1.4 km sec⁻¹, as there is very little difference between pixel-convolved Gaussians with standard deviations smaller than Δv . We restrict $v_{\text{los}}(x, y, i)$ to be within 45 km sec⁻¹ of the line-of-sight velocity at ℓ_x, b_y , and d_i corresponding to the IAU-recommended 220 km sec⁻¹ flat rotation curve and an 8.5 kpc separation between the Sun and the Galactic center. The middle panel of Figure 2 shows a cartoon solution to the \mathcal{L}_u optimization problem. Qualitative features of the cartoon such as the close match between the model and observed PPV 3-cubes, the magnitudes of deviations from the rotation curve, and the spatial coherence of deviations from the rotation curve are also typical of the actual solution to the unregularized problem.

This solution is not necessarily unique. While we have reduced the number of parameters in the problem by making assumptions about the structure of the ISM in PPDV space, there are still situations in which different $v_{\text{los}}(x, y, i)$ 3-cubes produce equivalent $m(x, y, j)$ 3-cubes. For example, if $m(x_0, y_0, i_0) = m(x_0, y_0, i_1)$, then $v_{\text{los}}(x_0, y_0, i_0)$ and $v_{\text{los}}(x_0, y_0, i_1)$ are interchangeable.

To deal with this problem, we introduce external information. The ℓ and b extent of the voxels in the 3-cubes and 4-cubes are often smaller than coherent structures such as molecular clouds. We may therefore expect voxels that share an ℓ or b boundary to have similar line-of-sight velocities. We encode this expectation into a regularized objective function, \mathcal{L}_r , by adding a term penalizing large differences between the line-of-sight velocities of voxels with shared ℓ or b boundaries:

$$\begin{aligned} \mathcal{L}_r(v_{\text{los}}(x, y, i), \sigma_v(x, y, i)) = & \frac{1}{\sigma_u^2} \mathcal{L}_u(v_{\text{los}}(x, y, i), \sigma_v(x, y, i)) + \\ & \frac{1}{2\sigma_r^2} \sum_{x, y} (v_{\text{los}}(x+1, y, i) - v_{\text{los}}(x, y, i))^2 + \\ & \frac{1}{2\sigma_r^2} \sum_{x, y} (v_{\text{los}}(x, y+1, i) - v_{\text{los}}(x, y, i))^2, \end{aligned} \quad (9)$$

where σ_u and σ_r are parameters that set the relative strengths of the model-observation residuals and regularization terms in driving the solution. We set σ_u to a value corresponding to 0.05 magnitudes of reddening. This is the standard deviation reported by GSF for the distribution of residuals between the GSF reddening 3-cube integrated along the distance axis and a Planck $\tau_{353\text{GHz}}$ -based integrated reddening map. We set σ_r to 5 km/s. This is approximately the value of the cloud-cloud dispersion found by Clemens (1985).

The bottom panel of Figure 2 shows a cartoon representation of a regularized solution. The regularization is represented by springs. As before, the qualitative features of the cartoon are realistic.

We evaluate the accuracy of the unregularized and regularized KT-derived v_{los} 3-cubes in the next section. For clarity, we will refer to the discretized 3-cubes as $v_{\text{los}}(\ell, b, d)$.

3.2. Checking the Kinetic Tomography solution

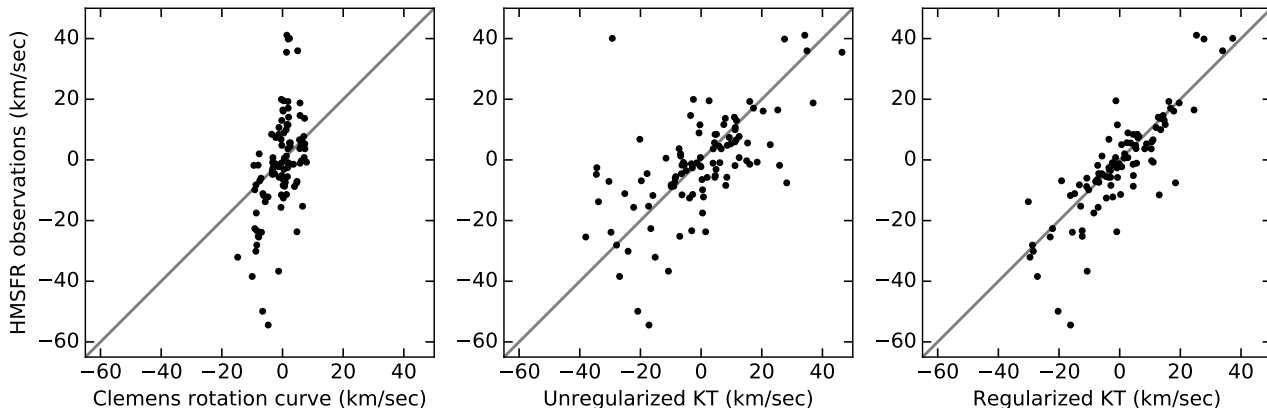


FIG. 3.— The three panels show the residuals from flat rotation of the Reid et al. (2014) HMSFR line-of-sight velocity observations on the y-axis and residuals from flat rotation of the values of the Clemens (1985) rotation curve (left panel), the unregularized kinetic tomography solution (middle panel), and the regularized kinetic tomography solution (right panel) at the positions of the HMSFRs on the x-axis. In each panel, the one-to-one line is indicated in gray.

The Reid14 HMSFRs (see §2.3) are embedded in, born from, and, presumably, moving with dense molecular gas. The v_{los} of an HMSFR should therefore be similar to the v_{los} of the ISM at the HMSFR’s location in PPD space. We can use this property to check the accuracy of our KT-derived $v_{\text{los}}(\ell, b, d)$ 3-cubes by comparing the HMSFRs’ observed v_{los} to the v_{los} KT assigns to the HMSFRs’ ℓ , b , and d values. In this section, we make this comparison for the $v_{\text{los}}(\ell, b, d)$ fields associated with a radially-varying rotation curve from Clemens (1985), unregularized KT, and regularized KT.

We must propagate the uncertainty on the distance to an HMSFR when comparing the HMSFR’s observed v_{los} to a $v_{\text{los}}(\ell, b, d)$ field. The typical uncertainty on the parallax of a Reid14 HMSFR is between 5 and 10% and is implicitly assumed to be Gaussian.

Following the discussion in Bovy et al. (2009), we assume that this Gaussian parallax uncertainty propagates linearly to a Gaussian distance uncertainty.

We assume that the uncertainty on the assigned $v_{\text{los}}(\ell, b, d)$ value is also approximately Gaussian. Consider an HMSFR, s . To compute the mean μ_s and standard deviation σ_s of this HMSFR’s distribution of possible $v_{\text{los}}(\ell, b, d)$ values, we start by drawing possible distance values d_t from $p_s(d)$. Here, $p_s(d)$ is the distribution over possible distances to HMSFR s and t is an index over draws. For each draw, we extract $v_{\text{los}}(\ell = \ell_s, b = b_s, d = d_t)$ from $v_{\text{los}}(\ell, b, d)$. The mean and standard deviation of these extracted line-of-sight velocity values are μ_s and σ_s .

In Figure 3, we show a comparison of the HMSFRs’ observed v_{los} values and the mean v_{los} values extracted from line-of-sight velocity fields corresponding to the Clemens (1985) rotation curve, unregularized KT, and regularized KT. To highlight the peculiar motions of the HMSFRs, these values are shown with the line-of-sight velocity corresponding to a flat 220 km/sec rotation and a Sun-Galactic center separation of 8.5 kpc subtracted off. The similarity of a velocity field to the HMSFR observations is indicated by how close points derived from

that velocity field tend to be to the one-to-one line in the appropriate panel of Figure 3. It should be clear from visual inspection that the radially-varying rotation curve does not accurately predict the velocities of the HMSFRs and that both versions of KT are clearly more accurate. There is also a significant, though possibly less visually obvious, improvement from unregularized to regularized KT.

To get a quantitative estimate of this improvement, we can compute the reduced χ^2 values of the two sets of velocity estimates. The reduced χ^2 value is given by the expression

$$\chi^2 = \frac{1}{\nu} \sum_s \left(\frac{v_{\text{los } s} - \mu_s}{\sigma_s} \right)^2, \quad (10)$$

where ν is the number of degrees of freedom in the problem and S is the number of observations. If the uncertainties are Gaussian and correctly estimated, the reduced χ^2 value should be approximately equal to 1. If we assume the regularization parameter counts against the number of degrees of freedom, the reduced χ^2 values of the unregularized and regularized KT solutions are 5 and 3, respectively.

The χ^2 value of the regularized KT solution is driven by 5 catastrophic outliers. If we remove these outliers, the reduced χ^2 values of the unregularized and regularized KT solutions drop to 4 and 1.3, respectively. We consider the advantage of regularized over unregularized KT to be sufficient to adopt the regularized KT solution as *the* KT solution, and will refer to it as such below.

At the positions of 94 of 99 HMSFRs, KT performs remarkably well.

3.3. Catastrophic outliers

In this section, we discuss the five HMSFRs where the KT-derived and directly observed v_{los} values are clearly inconsistent. The distribution of standardized residuals of the HMSFR sample is shown in panel (a) of Figure 4; the five outlier HMSFRs are marked in red. The reason

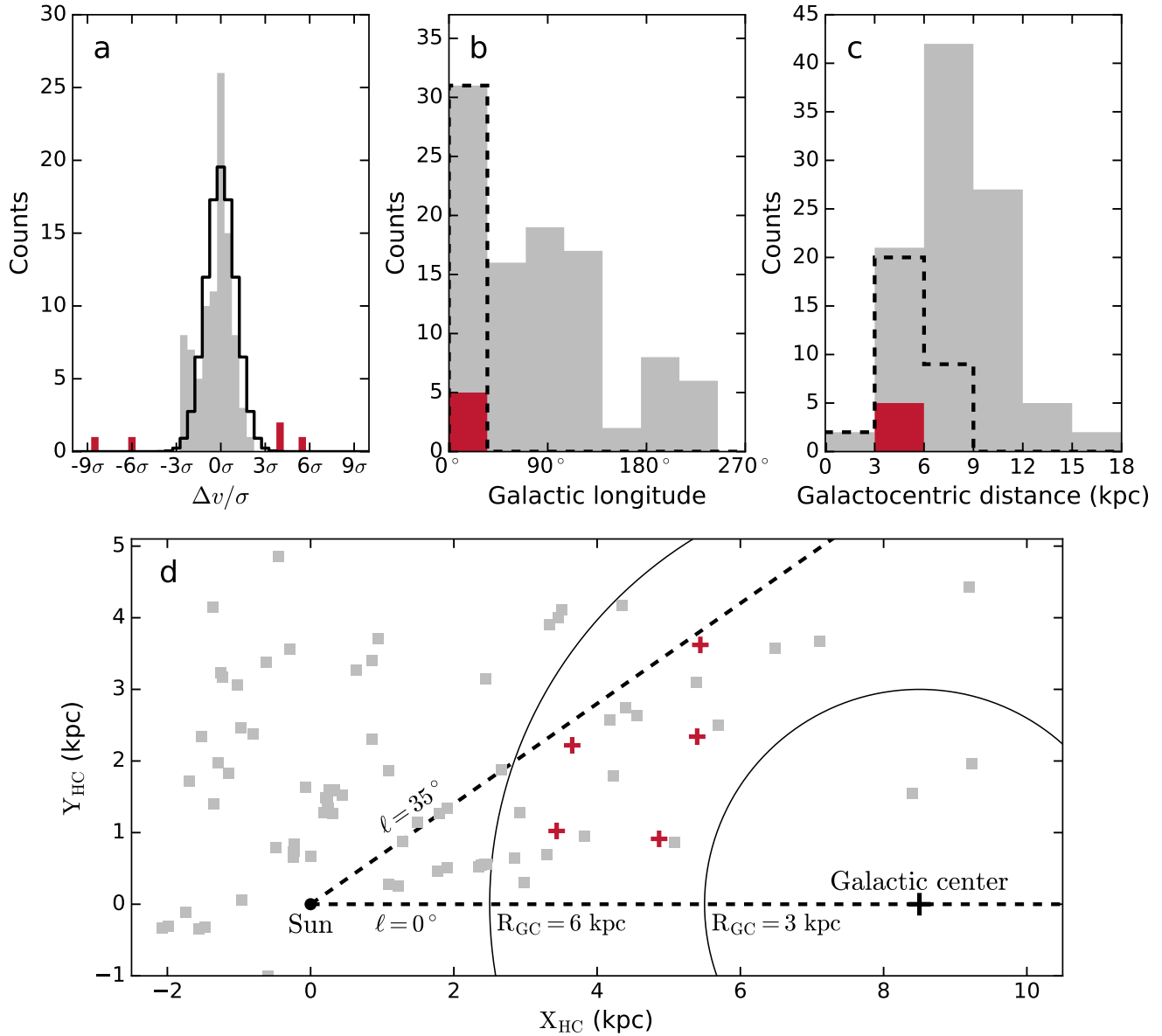


FIG. 4.— (a) The distribution of residuals per standard deviation (*standardized* residuals) between the observed and KT-derived line-of-sight velocities of HMSFRs (filled histogram), with an appropriately-scaled standard normal distribution for reference (empty histogram). We consider HMSFRs with $|\Delta v/\sigma| > 3$ to be outliers. In all three panels, outliers are denoted in red and non-outliers are denoted in gray. (b) The distribution of the HMSFRs' Galactic longitudes. HMSFRs with $\ell < 35^\circ$ are outlined with a dashed black line. (c) The distribution of the HMSFRs' Galactocentric radii. HMSFRs with $\ell < 35^\circ$ are outlined with a dashed black line. (d) The locations of the HMSFRs in heliocentric Cartesian coordinates, where X_{HC} increases towards $\ell = 0^\circ$ and Y_{HC} increases towards $\ell = 90^\circ$.

why these five HMSFRs, specifically, are outliers seems to be that they are located in a particularly complex and confused part of the Galaxy. From panels (b), (c), and (d) of the same Figure, one can see that all of the outliers are in the inner Galaxy and relatively close to the Galactic center, at $0^\circ < \ell < 35^\circ$ and $R_{\text{GC}} < 6$ kpc. We note that four of these five outliers are in the area of sky covered by the GASS Galactic HI survey. We doubt GASS is to blame as it has relatively mild systematic errors and is intermediate in spatial resolution among the

three surveys we use. If we compare the distributions of all HMSFRs and outlier HMSFRs in ℓ and R_{GC} , it should be apparent that the distribution of outliers does not follow the distribution of all HMSFRs.

We can quantify this deviation by assuming that any HMSFR is equally likely to be one of the five outliers and computing the probability of finding all five in a randomly chosen subsample of the same size as one of our two selections. Since a single HMSFR can not be present in a sample twice, the appropriate distribution to use for

this calculation is the hypergeometric distribution, which assumes that the subsample is chosen without replacement. The probability of finding all five outliers in a random subsample the size of either of the two selections is, in both cases, much less than 1%.

The implication of this difference in distribution is that there is an actual enhancement in the rate at which KT produces incorrect solutions for sightlines that pass near the Galactic center. There are three main reasons why KT would be less accurate along a sightline at low ℓ . Firstly, these sightlines pass through more matter, due to both the increase in ISM surface density at low R_{GC} and the fact that geometrically, a sightline at low ℓ will pass through more of the far side of the Galaxy. Having more matter along a sightline increases the complexity of the problem — more parcels of ISM in PPD space have to be correctly associated with velocity components in PPV space. Secondly, due in part to the increase in ISM surface density with decreasing R_{GC} , the well-measured part of the PPD cube does not extend as far towards the inner Galaxy as it does towards the outer Galaxy (GSF). If the PPD cube is not accurate past some distance, then we are unlikely to obtain a correct solution past that distance. Thirdly, the Galactic bar can stir interstellar matter up, inducing a complicated velocity field with lots of structure on small spatial scales. Given that the distance extent of PPD voxels near the inner Galaxy is of order a kpc, we could simply have insufficient spatial resolution to map this bar-induced velocity field.

Two of these three reasons are caused by being at low ℓ while the other is caused by being in a specific R_{GC} range. To try to differentiate between the two possible causes, we can repeat the quantitative exercise described above with the R_{GC} selection as the subsample and the ℓ selection, rather than the full set of 99 HMSFRs, as the population. The probability of finding all five outliers in a random 21-element subsample of the 31-element ℓ selection is about 11%. This probability is low but not unreasonable; we cannot differentiate between the two causes in this way.

4. DISCUSSION

4.1. Qualitative features of the solution

We have established that KT is, quantitatively-speaking, correct at the locations of 94 dense, star-forming knots of interstellar matter and we have found a heuristic explanation for why KT is quite incorrect at the locations of 5 other dense, star-forming knots of interstellar matter. Given the lack of other published $v_{los}(\ell, b, d)$ measurements, we cannot quantitatively establish that our $v_{los}(\ell, b, d)$ 3-cube is also correct away from these 99 locations. There is, however, the previously published two-dimensional peculiar velocity map of Brand & Blitz (1993).

A two-dimensional peculiar velocity field derived from the KT solution is shown in Figure 5. This peculiar velocity field is derived by subtracting the line-of-sight velocity field corresponding to a flat rotation curve from the $v_{los}(\ell, b, d)$ 3-cube and taking the mass-weighted average of this peculiar velocity 3-cube along the b axis over the range $-2.5^\circ \leq b \leq +2.5^\circ$. The peculiar velocity field of Brand & Blitz (1993) is shown in Figure 10 of their paper. Qualitatively, the two peculiar velocity fields are

quite similar. The locations, extents, and signs of regions of coherent peculiar velocity are, for the most part, the same in both maps. There are, however, some differences in the magnitudes of peculiar velocities in these regions.

We can also qualitatively compare the typical scale of coherent velocity fluctuations in our two-dimensional peculiar velocity field to measurements of this scale from the literature. Much of the power in the peculiar velocity field of disc gas (e.g. Clemens 1985) and disc stars (e.g. Bovy et al. 2015) is found on scales of about 2 kpc. This is consistent, at least by eye, with the typical extent of a coherent peculiar velocity fluctuation in our two-dimensional peculiar velocity field.

Figure 5 also shows the line-of-sight velocities at the ℓ , b , and d of HMSFRs according to our $v_{los}(\ell, b, d)$ 3-cube and the actual observations. All of the HMSFRs are in b range we averaged over to produce our two-dimensional peculiar velocity field, but there are clear differences between the observed HMSFR line-of-sight velocities and the two-dimensional velocity field. If we assume the two-dimensional velocity field is, for the most part, correct, then these differences imply peculiar velocity fluctuations on spatial scales of order the typical height off the plane of an HMSFR, or 10-100 pc. The same differences, in most cases, are also seen in the velocities assigned based on our $v_{los}(\ell, b, d)$ 3-cube. Combining these facts, we can heuristically conclude that we are more-or-less correctly detecting peculiar velocity fluctuations on both large, multi-kpc scales and small, 100 pc scales.

4.2. Unmet assumptions and their potential consequences

Our treatment of the data, our parametrization of the 4-dimensional structure of the ISM, and our regularization scheme are all based on assumptions that are not always met. We convert our distributions of reddening, HI emission, and CO emission to distributions of absolute matter content in PPV and PPD space assuming there is a single, linear function relating the amount of each tracer to an amount of matter. Due to variations in the dust-to-gas and CO-to-H₂ ratios and radiative transfer effects such as self-absorption, these functions are neither uniquely defined nor linear. We have implicitly assumed that the PPD and PPV 3-cubes are projections of the same part of the same PPDV 4-cube. This is not the case — the PPV 3-cube is an integral of the PPDV 4-cube to an effectively infinite distance while the most distant well-measured voxels in the PPD 3-cube are only ~ 10 kpc away from the Sun (GSF). Our parametrization of the 4-dimensional structure of the ISM assumes that the ISM's velocity distribution within a PPD voxel can be described by a mean, which is close to the value of the Galactic rotation curve at the voxel's center, and a dispersion. This will not be true for a voxel that contains a shock or is large compared to the spatial scale of velocity fluctuations; because our distance resolution is constant in log space, this effect will apply to all voxels beyond some distance. Our regularization scheme assumes that the velocity distributions of voxels with shared ℓ or b boundaries will be similar. This will once again not be true of voxels that cross shocks or are sufficiently large.

Despite the fact that its underlying assumptions do not hold over some of the solution domain, KT produces a $v_{los}(\ell, b, d)$ solution that quantitatively agrees with in-

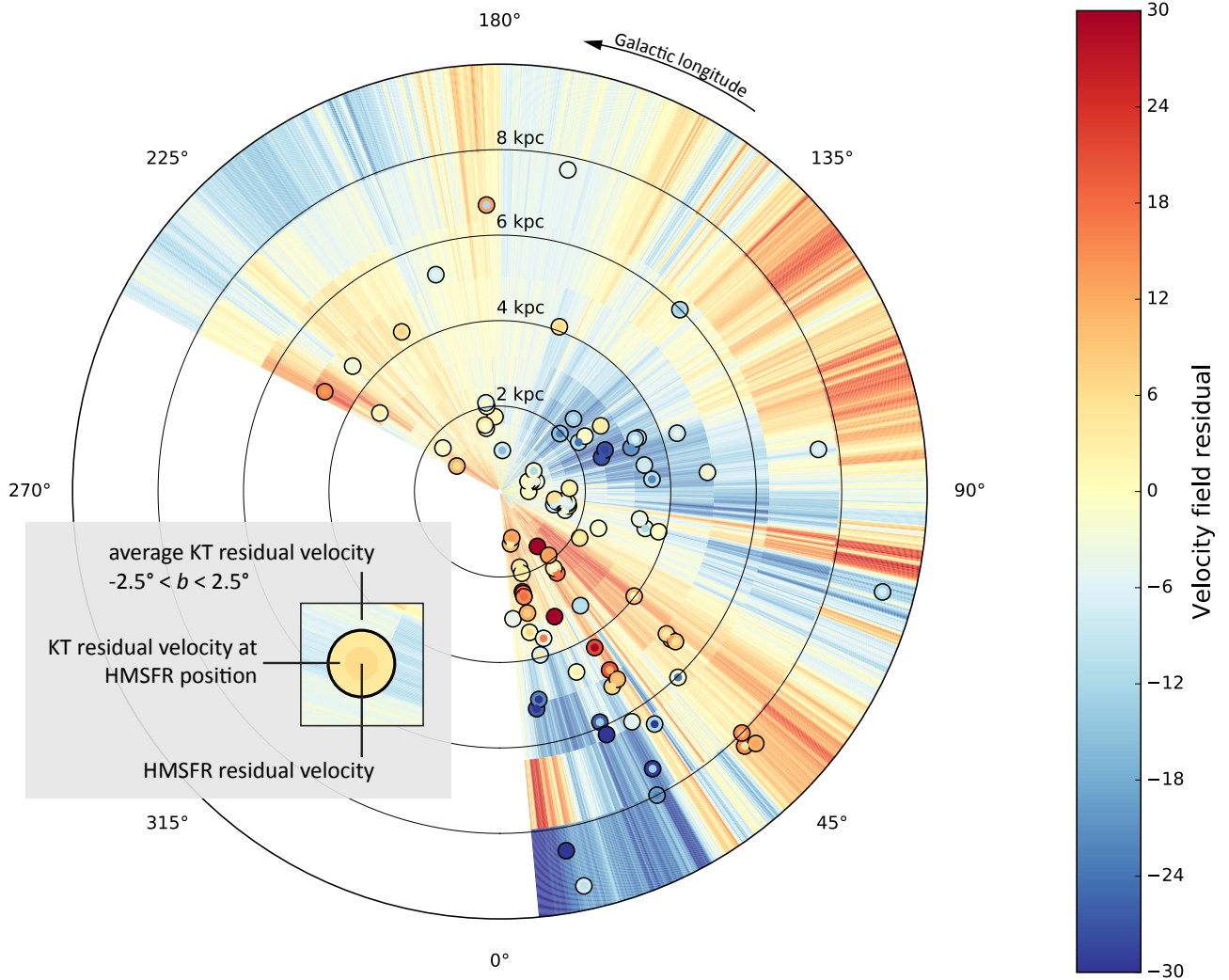


FIG. 5.— Colors show the difference between an observed or estimated v_{los} and the value predicted from a flat rotation curve. The background is an average of the kinetic tomography-derived velocity field over $-2.5^\circ \leq b \leq +2.5^\circ$. The color of the inner part of each circle is the v_{los} of an HMSFR. The color of the outer ring of each circle is the value of the kinetic tomography-derived velocity field at the position (including b) of the HMSFR.

dependent $v_{\text{los}}(\ell, b, d)$ observations (see §3.2). We can resolve this conflict by concluding that (1) the assumptions, as stated, are too strict or by arguing that (2) the particular set of $v_{\text{los}}(\ell, b, d)$ observations are not representative of the ISM as a whole.

The fact that all of the poorly-reproduced HMSFRs lie in the inner Galaxy (see §3.3), where our assumptions tend to not be met more often and more egregiously than in the outer Galaxy, lends credence to the resolution (1). Perhaps KT is robust to some level of its assumptions not being met and this level is only exceeded in some parts of the inner Galaxy. As an example of the extent to which the inner Galaxy does not meet our assumptions, consider the fact that the maximum distance from the Sun to which the PPD 3-cube is accurate in the inner Galaxy is approximately 5 kpc GSF while the PPV 3-cube is integrated out to the far edge of the Galaxy.

Resolution (2), that the comparison observations are

atypical, seems unlikely considering the qualitative structure of the solution (see §4.1) but cannot yet be ruled out empirically. The comparison observations are of HMSFRs, which by definition will be associated with large overdensities of molecular gas. It is possible that the check we have performed in §3.2 only applies to dense molecular gas. For example, KT could merely be finding the most massive object along a sightline through the PPD 3-cube and associating it with the most massive objective in the corresponding sightline through the PPV 3-cube. This would usually correctly assign a distance and velocity to something like a GMC but would fail for less concentrated neutral gas. However, it would be difficult to reconcile the clear and non-trivial larger-scale velocity structure seen in the previous section with a scenario in which the KT solution is only correct at extreme overdensities.

To get a more quantitative and broadly applicable un-

derstanding of when KT works and fails, we would need either a less density-biased set of independent $v_{\text{los}}(\ell, b, d)$ measurements or a set of artificial injection tests. These artificial injection tests would consist of numerical experiments in which we artificially observe a model galaxy's ISM, reconstruct the v_{los} field from these artificial observations using KT, and compare the reconstructed and input model v_{los} fields. To the best of our knowledge, there are no currently available catalogs of these sorts of less density-biased measurements, ruling out option one. The steps involved in artificial injection tests, particularly simulating galaxies at sufficiently high resolution and producing artificial observations in a way that includes the non-trivial systematics in the actual observations, are complicated enough to put option two beyond the scope of this work.

Both of these options are plausible directions for future work. The $1.527 \mu\text{m}$ diffuse interstellar band (DIB), for instance, has been mapped over much of the northern sky by the APOGEE survey (Zasowski et al. 2015). Observations of this DIB towards APOGEE stars with known distances could potentially be used to build catalogs or even maps of $v_{\text{los}}(\ell, b, d)$ using an independent dataset. Artificial injection tests are conceptually straightforward, though they do require a substantial investment of time and computational resources. While neither option is easy, we would argue that some combination of the two will be necessary before the KT-derived v_{los} 3-cube can be trusted away from the HMSFRs of Reid14.

5. CONCLUSION

In this work we developed a method for measuring the radial velocity of parcels of the interstellar medium of a measured distance, which we dubbed Kinetic Tomography. We argued that this method is important as a tool for measuring converging and diverging flows around our Galaxy as well as for detecting large scale deviations

from assumed rotation curves. The method takes as inputs the three-dimensional distribution of dust in our Galaxy measured from stellar photometry and the emission spectra from Galactic CO and HI. We developed a technique that assigns each 3D parcel of ISM from the dust map a line-of-sight velocity and line-of-sight velocity width, in order to best reproduce the observed CO and HI data. We found that we can improve the fidelity of our solution by implementing Tikhonov regularization, effectively coupling the line-of-sight velocity of adjacent pixels.

As a test of our method we compare our results to independent measurements of HMSFRs from Reid14, which contain both distances and line-of-sight velocity information. We find that of the 99 HMSFRs in the area of sky we study, 94 are consistent with our results and 5 are outliers, all of which lie near Galactic center. This consistency indicates our map is an accurate representation of the velocity field of the ISM, at least in denser regions consistent with HMSFRs. We also find qualitative consistency with the peculiar velocity maps of Brand & Blitz (1993).

Here we conclude that KT can be a very powerful tool for the study of the velocity structure of the Galactic ISM. In future work, we will investigate what KT can tell us about the Galactic rotation curve, streaming motions within the Galactic disk, and the vertical structure of Galactic flows.

6. ACKNOWLEDGEMENTS

We thank the anonymous referee for their concise and constructive comments. We thank Greg Green and Edie Schlafly for guidance on the use of the GSF reddening cube. This material is based upon work supported by the National Science Foundation under Grant No. 1616177 and the Space Telescope Science Institute Director's Discretionary Fund.

REFERENCES

- Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, *ARA&A*, 51, 207
- Bonnell, I. A., Dobbs, C. L., Robitaille, T. P., & Pringle, J. E. 2006, *MNRAS*, 365, 37
- Bovy, J., Bird, J. C., García Pérez, A. E., et al. 2015, *ApJ*, 800, 83
- Bovy, J., Hogg, D. W., & Rix, H.-W. 2009, *ApJ*, 704, 1704
- Brand, J., & Blitz, L. 1993, *A&A*, 275, 67
- Carroll-Nellenback, J. J., Frank, A., & Heitsch, F. 2014, *ApJ*, 790, 37
- Clark, P. C., Glover, S. C. O., Klessen, R. S., & Bonnell, I. A. 2012, *MNRAS*, 424, 2599
- Clemens, D. P. 1985, *ApJ*, 295, 422
- Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, *ApJ*, 547, 792
- Dobbs, C. L. 2008, *MNRAS*, 391, 844
- Dobbs, C. L., Pringle, J. E., & Burkert, A. 2012, *MNRAS*, 425, 2157
- Euler, L. 1757, *Académie R. Sci. Belles-Lett. Berlin Mém.*, 11
- Foster, T., & MacWilliams, J. 2006, *ApJ*, 644, 214
- Fujimoto, Y., Tasker, E. J., Wakayama, M., & Habe, A. 2014, *MNRAS*, 439, 936
- Goldsmith, P. F., Li, D., & Krčo, M. 2007, *ApJ*, 654, 273
- Green, G. M., Schlafly, E. F., Finkbeiner, D. P., et al. 2015, *ApJ*, 810, 25
- Heitsch, F., Slyz, A. D., Devriendt, J. E. G., Hartmann, L. W., & Burkert, A. 2006, *ApJ*, 648, 1052
- Jenkins, E. B. 2009, *ApJ*, 700, 1299
- Kalberla, P. M. W., Burton, W. B., Hartmann, D., et al. 2005, *A&A*, 440, 775
- Kalberla, P. M. W., McClure-Griffiths, N. M., Pisano, D. J., et al. 2010, *A&A*, 521, A17
- Kim, W.-T., & Ostriker, E. C. 2002, *ApJ*, 570, 132
- Lallement, R., Vergely, J.-L., Valette, B., et al. 2014, *A&A*, 561, A91
- Levine, E. S. 2006, *Sci*, 312, 1773
- Marshall, D. J., Robin, A. C., Reylé, C., Schultheis, M., & Picaud, S. 2006, *A&A*, 453, 635
- Peek, J. E. G. 2013, *ApJ*, 766, L6
- Peek, J. E. G., Heiles, C., Douglas, K. A., et al. 2011, *ApJS*, 194, 20
- Planck Collaboration and Fermi Collaboration, Ade, P. A. R., Aghanim, N., et al. 2015, *A&A*, 582, A31
- Reid, M. J., Dame, T. M., Menten, K. M., & Brunthaler, A. 2016, *ApJ*, 823, 77
- Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2014, *ApJ*, 783, 130
- Roberts, W. W. J. 1972, *ApJ*, 173, 259
- Roberts, W. W. J., & Stewart, G. R. 1987, *ApJ*, 314, 10
- Savage, B. D., & Sembach, K. R. 1996, *ARA&A*, 34, 279
- Shane, W. W. 1972, *A&A*, 16, 118
- Shetty, R., Vogel, S. N., Ostriker, E. C., & Teuben, P. J. 2007, *ApJ*, 665, 1138
- Tasker, E. J., & Tan, J. C. 2009, *ApJ*, 700, 358
- Vazquez Semadeni, E., Gomez, G. C., Jappsen, A. K., et al. 2007, *ApJ*, 657, 870
- Visser, H. C. D. 1980a, *A&A*, 88, 159
- . 1980b, *A&A*, 88, 149

Welsh, B. Y., Lallement, R., Vergely, J.-L., & Raimond, S. 2010, A&A, 510, 54

Zasowski, G. 2015, MmSAI, 86, 521

Zasowski, G., Menard, B., Bizyaev, D., et al. 2015, ApJ, 798, 35